

# Second-Order Stochastic Dominance Constraints Induced by Mixed-Integer Linear Recourse

Ralf Gollmer, Uwe Gotzes, Rüdiger Schultz\*

Department of Mathematics

University of Duisburg-Essen, Campus Duisburg

Lotharstr. 65, D-47048 Duisburg, Germany

## Abstract

We introduce stochastic integer programs with dominance constraints induced by mixed-integer linear recourse. Closedness of the constraint set mapping with respect to perturbations of the underlying probability measure is derived. For discrete probability measures, large-scale, block-structured, mixed-integer linear programming equivalents to the dominance constrained stochastic programs are identified. For these models, a decomposition algorithm is proposed. Computational tests with instances from power optimization and Sudoku puzzling conclude the paper.

**Key Words.** Stochastic integer programming, stochastic dominance, mixed-integer optimization.

**AMS subject classifications.** 90C15, 90C11, 60E15.

## 1 Introduction

In recent years, stochastic dominance has gained some attraction in stochastic programming. In [10, 11, 12, 22] the authors have introduced stochastic programs with dominance constraints of different order involving generic random variables. They have studied basic structural properties of these models and proposed algorithms for their solution.

Notions of stochastic dominance introduce partial orders on families of random variables. When preferring small outcomes to big ones, a (real-valued) random variable  $X$  is said to dominate a random variable  $Y$  to first order ( $X \succeq_1 Y$ ) iff  $\mathbb{E}h(X) \leq \mathbb{E}h(Y)$  for all nondecreasing functions  $h$  for which both expectations exist.  $X$  is said to dominate  $Y$  to second order ( $X \succeq_2 Y$ ) iff

$$\mathbb{E}h(X) \leq \mathbb{E}h(Y) \quad (1)$$

for all nondecreasing convex functions  $h$  for which both expectations exist. Dominance of first order implies that of second order to hold, but not vice versa. We refer to [21] and the references therein for background on stochastic dominance.

The objects of interest in the present paper are stochastic programs with second-order dominance constraints involving random variables that result from the dynamics met in two-stage stochastic programming. To be more specific, let us consider the following random mixed-integer linear program

$$\min\{c^\top x + q^\top y : Tx + Wy = z(\omega), x \in X, y \in \mathbb{Z}_+^m \times \mathbb{R}_+^{m'}\} \quad (2)$$

together with the information constraint that, in the first stage,  $x$  must be selected prior to observing  $z(\omega)$ , and afterwards, in a second stage,  $y$  has to be selected. This condition often is referred to as nonanticipativity of  $x$ . We assume that the vectors and matrices in (2) have conformable dimensions, that  $W$  has rational entries, and that  $X \subseteq \mathbb{R}^m$  is a nonempty polyhedron, possibly involving integer requirements to components of  $x$ .

---

\*schultz@math.uni-duisburg.de

In traditional two-stage stochastic programming, see [6, 18, 23, 26], the aim is to optimize first-stage decisions. To this end, well-defined optimization problems in  $x$ , often called deterministic equivalents, are formulated. The principal construction is as follows. Rewrite (2) as

$$\begin{aligned} \min_x \left\{ c^\top x + \min_y \{ q^\top y : Wy = z(\omega) - Tx, y \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'} \} : x \in X \right\} \\ = \min_x \{ c^\top x + \Phi(z(\omega) - Tx) : x \in X \} \end{aligned} \quad (3)$$

where

$$\Phi(t) := \min \{ q^\top y : Wy = t, y \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'} \}. \quad (4)$$

One possibility to look at (3) is to recognise a family of random variables

$$\left( c^\top x + \Phi(z(\omega) - Tx) \right)_{x \in X}, \quad (5)$$

and to understand (3) as the problem of finding a “best” member in this family. The most straightforward way to make this selection is to compare the random variables by their expectations, leading to the deterministic equivalent

$$\min \{ \mathbb{E}[c^\top x + \Phi(z(\omega) - Tx)] : x \in X \}. \quad (6)$$

In the literature this problem is known as the (classical) two-stage stochastic program with (mixed-integer) linear recourse, [6, 18, 23, 26]. The risk neutral problem (6) can be extended to a model involving risk aversion if the random variables  $f(x, \omega) := c^\top x + \Phi(z(\omega) - Tx)$  get compared via statistical parameters reflecting risk. With  $\mathcal{R}$  a generic such parameter and  $\rho > 0$  a fixed weight factor, the mean-risk extension of (6) reads

$$\min \{ \mathbb{E}[f(x, \omega)] + \rho \cdot \mathcal{R}[f(x, \omega)] : x \in X \}.$$

Risk measures  $\mathcal{R}$  that were used in this context include both quantile-based (excess probability, value-at-risk, conditional value-at-risk) and deviation-based measures (expected excess, semideviation), see [1, 13, 19, 20, 28, 29].

The starting point of the investigations in the present paper is to identify “acceptable” members of (5) rather than looking for a “best” among them. We assume that a random benchmark  $a(\omega)$  is given that reflects an acceptance threshold for the costs  $f(x, \omega)$  resulting from the two-stage dynamics in (2). We will consider  $x \in X$  acceptable iff  $f(x, \omega) \succeq_2 a(\omega)$ . Over all acceptable  $x$  we minimize an objective function  $g : \mathbb{R}^m \rightarrow \mathbb{R}$ . This leads to the following stochastic program with second-order dominance constraints induced by mixed-integer recourse

$$\min \{ g(x) : f(x, \omega) \succeq_2 a(\omega), x \in X \}. \quad (7)$$

This model, while interesting in its own, is closely related to the counterpart model where second-order dominance is replaced by first-order dominance, see [14] for an analysis of the latter. Since second-order dominance is the weaker notion, (7) is a relaxation of the first-order model, see [22] for related work.

Our paper is organized as follows. In Section 2 we study some structural properties of (7). Section 3 is devoted to algorithmic considerations, and in the final section we report some computational results.

## 2 Structural Properties

The aim of this section is to provide a framework such that the objects in (7) are well-defined, and to derive some basic structural properties of (7).

Let us come back to the definition of second-order stochastic dominance. It is well-known, see [21] for a proof, that (1) is already valid if it holds for all “wedge” functions  $h(\cdot) := \max\{\cdot - \eta, 0\} =: [\cdot - \eta]_+$ ,  $\eta \in \mathbb{R}$ . Let  $\mathcal{P}(\mathbb{R}^s), \mathcal{P}(\mathbb{R})$  be the sets of all Borel probability measures on  $\mathbb{R}^s$  and  $\mathbb{R}$ , and let  $\mu \in \mathcal{P}(\mathbb{R}^s)$  and  $\nu \in \mathcal{P}(\mathbb{R})$  denote the image measures induced by  $z(\omega)$ ,  $a(\omega)$  on  $\mathbb{R}^s$  and  $\mathbb{R}$ , respectively. The constraint  $f(x, \omega) \succeq_2 a(\omega)$  now can be equivalently expressed as

$$\int_{\mathbb{R}^s} [f(x, z) - \eta]_+ \mu(dz) \leq \int_{\mathbb{R}} [a - \eta]_+ \nu(da) \quad \forall \eta \in \mathbb{R}, \quad (8)$$

provided all objects in (8) are well-defined. Let us start with  $f(x, z) := c^\top x + \Phi(z - Tx)$ . Recall that  $\Phi$  is the value function of a mixed-integer linear program, cf. (4). Assume

- (A1) (complete recourse)  $W(\mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}) = \mathbb{R}^s$ ,  
(A2) (sufficiently expensive recourse)  $\{u \in \mathbb{R}^s : W^\top u \leq q\} \neq \emptyset$ .

Then it holds that (A1) and (A2), together with the rationality of  $W$ , which was imposed as a basic assumption in the very beginning, imply that  $\Phi$  is real-valued and lower semicontinuous on  $\mathbb{R}^s$ , i.e.,  $\liminf_{t_n \rightarrow t} \Phi(t_n) \geq \Phi(t)$  for all  $t \in \mathbb{R}^s$ , [3, 7]. Moreover, there exist  $\alpha > 0$ ,  $\beta > 0$  such that for all  $t_1, t_2 \in \mathbb{R}^s$

$$|\Phi(t_1) - \Phi(t_2)| \leq \alpha \|t_1 - t_2\| + \beta. \quad (9)$$

Without integer requirements in the second stage (linear recourse), linear programming duality, together with (A1) and (A2), imply

$$\Phi(t) = \min\{q^\top y : Wy = t, y \geq 0\} = \max\{t^\top u : W^\top u \leq q\} = \max_{\ell=1, \dots, L} d_\ell^\top t \quad (10)$$

where  $d_\ell, \ell = 1, \dots, L$ , are the vertices of  $\{u : W^\top u \leq q\}$ . Hence,  $\Phi$  is piecewise linear and convex in this case.

This settles well-posedness of the integrands in (8). For finiteness of the integrals we assume

$$(A3) \quad (\text{finite first moments}) \quad \int_{\mathbb{R}^s} \|z\| \mu(dz) < \infty, \quad \int_{\mathbb{R}} |a| \nu(da) < \infty.$$

Using (9) and the fact that (A2) implies  $\Phi(0) = 0$ , we obtain that for fixed  $x$  there is a constant  $\kappa > 0$  such that

$$|[f(x, z) - \eta]_+| \leq \alpha \|z\| + \kappa \quad \forall z \in \mathbb{R}^s.$$

Hence, (A1)-(A3) imply that the integral on the left in (8) is always finite. For the integral on the right, (A3) ensures this property.

In accordance with (A3) we denote by  $\mathcal{P}_1(\mathbb{R}^s), \mathcal{P}_1(\mathbb{R})$  the subsets of  $\mathcal{P}(\mathbb{R}^s), \mathcal{P}(\mathbb{R})$  with measures having finite first moment. We fix  $\nu \in \mathcal{P}_1(\mathbb{R})$  and consider the multifunction  $C : \mathcal{P}_1(\mathbb{R}^s) \rightarrow 2^{\mathbb{R}^m}$  where

$$C(\mu) := \{x \in \mathbb{R}^m : f(x, z) \succeq_2 a, x \in X\}. \quad (11)$$

The space  $\mathcal{P}_1(\mathbb{R}^s)$  is equipped with weak convergence of probability measures ([4]). A sequence  $\{\mu_n\}$  in  $\mathcal{P}_1(\mathbb{R}^s)$  is said to converge weakly to  $\mu \in \mathcal{P}_1(\mathbb{R}^s)$ , written  $\mu_n \xrightarrow{w} \mu$ , if for any bounded continuous function  $h : \mathbb{R}^s \rightarrow \mathbb{R}$  it holds  $\int_{\mathbb{R}^s} h(z) \mu_n(dz) \rightarrow \int_{\mathbb{R}^s} h(z) \mu(dz)$  as  $n \rightarrow \infty$ .

Our aim is to show that  $C$  is a closed multifunction on  $\mathcal{P}_1(\mathbb{R}^s)$ . This means that for arbitrary  $\mu \in \mathcal{P}_1(\mathbb{R}^s)$  and sequences  $\mu_n \in \mathcal{P}_1(\mathbb{R}^s), x_n \in C(\mu_n)$  with  $\mu_n \xrightarrow{w} \mu$  and  $x_n \rightarrow x$  it follows that  $x \in C(\mu)$ .

**Lemma 2.1** *Let  $\mu_n, \mu \in \mathcal{P}(\mathbb{R}^s)$  with  $\mu_n \xrightarrow{w} \mu$  and  $h : \mathbb{R}^s \rightarrow \mathbb{R}$  be lower semicontinuous with  $h(z) \geq 0 \forall z \in \mathbb{R}^s$ . Then*

$$\int_{\mathbb{R}^s} h(z) \mu(dz) \leq \liminf_n \int_{\mathbb{R}^s} h(z) \mu_n(dz).$$

**Proof:** We start with the bounded case and assume there exist  $\underline{h}, \bar{h} \in \mathbb{R}$  such that  $\underline{h} < h(z) < \bar{h} \forall z \in \mathbb{R}^s$ . Without loss of generality we assume  $0 < h(z) < 1 \forall z \in \mathbb{R}^s$  which can be achieved by affine scaling.

Fix  $k \in \mathbb{N}$  and consider the sets  $H_i := \{z \in \mathbb{R}^s : i/k < h(z)\}$ ,  $i = 0, \dots, k$ . Since  $h$  is lower semicontinuous,  $H_i$  is open for all  $i$ . It holds

$$\sum_{i=1}^k \frac{i-1}{k} \mu\left[\left\{z : \frac{i-1}{k} < h(z) \leq \frac{i}{k}\right\}\right] \leq \int_{\mathbb{R}^s} h(z) \mu(dz) \leq \sum_{i=1}^k \frac{i}{k} \mu\left[\left\{z : \frac{i-1}{k} < h(z) \leq \frac{i}{k}\right\}\right].$$

The sum on the right equals

$$\sum_{i=1}^k \frac{i}{k} (\mu[H_{i-1}] - \mu[H_i]) = \frac{1}{k} + \frac{1}{k} \sum_{i=1}^k \mu[H_i],$$

while the sum on the left is identical with

$$\sum_{i=1}^k \frac{i-1}{k} (\mu[H_{i-1}] - \mu[H_i]) = \frac{1}{k} \sum_{i=1}^k \mu[H_i].$$

Putting this together yields

$$\frac{1}{k} \sum_{i=1}^k \mu[H_i] \leq \int_{\mathbb{R}^s} h(z) \mu(dz) \leq \frac{1}{k} + \frac{1}{k} \sum_{i=1}^k \mu[H_i]. \quad (12)$$

By the Portmanteau Theorem (see [4], Theorem 2.1, p 11/12) we have for all  $i$

$$\mu[H_i] \leq \liminf_n \mu_n[H_i]. \quad (13)$$

Applying the left inequality in (12) to  $\mu_n$  and taking the limes inferior provides

$$\frac{1}{k} \liminf_n \sum_{i=1}^k \mu_n[H_i] \leq \liminf_n \int_{\mathbb{R}^s} h(z) \mu_n(dz),$$

and, together with (13),

$$\frac{1}{k} \sum_{i=1}^k \mu[H_i] \leq \liminf_n \int_{\mathbb{R}^s} h(z) \mu_n(dz).$$

Now we apply the right inequality in (12) and obtain

$$-\frac{1}{k} + \int_{\mathbb{R}^s} h(z) \mu(dz) \leq \liminf_n \int_{\mathbb{R}^s} h(z) \mu_n(dz).$$

With  $k \rightarrow \infty$  this yields the assertion for bounded  $h$ . For extension to unbounded non-negative  $h$  let  $r \in \mathbb{R}_+$  and consider the truncated function  $h_r : \mathbb{R}^s \rightarrow \mathbb{R}$  with

$$h_r(z) := \begin{cases} h(z) & , \text{ if } h(z) \leq r \\ r & , \text{ otherwise.} \end{cases}$$

Lower semicontinuity of  $h$  implies lower semicontinuity of  $h_r$  for all  $r \in \mathbb{R}_+$ . The assertion then is valid for  $h_r$ , since  $h_r$  is bounded. Moreover,  $h_r(z) \leq h(z) \forall z \in \mathbb{R}^s$ . This yields

$$\int_{\mathbb{R}^s} h_r(z) \mu(dz) \leq \liminf_n \int_{\mathbb{R}^s} h_r(z) \mu_n(dz) \leq \liminf_n \int_{\mathbb{R}^s} h(z) \mu_n(dz) \quad \forall r \in \mathbb{R}_+. \quad (14)$$

The Monotone Convergence Theorem (see for instance [5], Theorem 16.2, p. 211) yields

$$\int_{\mathbb{R}^s} h_r(z) \mu(dz) \longrightarrow \int_{\mathbb{R}^s} h(z) \mu(dz) \quad \text{for } r \rightarrow \infty.$$

Together with (14) this implies

$$\int_{\mathbb{R}^s} h(z) \mu(dz) \leq \liminf_n \int_{\mathbb{R}^s} h(z) \mu_n(dz),$$

and the proof is complete.  $\square$

**Proposition 2.2** Assume (A1)-(A3). Then the multifunction  $C$ , as defined in (11), is closed on  $\mathcal{P}_1(\mathbb{R}^s)$ .

**Proof:** Let  $\mu_n, \mu \in \mathcal{P}_1(\mathbb{R}^s)$  and  $x_n \in C(\mu_n)$  such that  $\mu_n \xrightarrow{w} \mu$  and  $x_n \rightarrow x$ . Closedness of  $X$  then immediately yields  $x \in X$ . According to (8),  $x_n \in C(\mu_n)$  implies

$$\int_{\mathbb{R}^s} [f(x_n, z) - \eta]_+ \mu_n(dz) \leq \int_{\mathbb{R}} [a - \eta]_+ \nu(da) \quad \forall \eta \in \mathbb{R}. \quad (15)$$

Notice that the integrands  $[f(\cdot, \cdot) - \eta]_+$  are non-negative and lower semicontinuous for all  $\eta \in \mathbb{R}$ . Together with Fatou's Lemma (see for instance [5], Theorem 16.3, p. 212), this implies

$$\begin{aligned} \int_{\mathbb{R}^s} [f(x, z) - \eta]_+ \mu(dz) &\leq \int_{\mathbb{R}^s} \liminf_k [f(x_k, z) - \eta]_+ \mu_n(dz) \\ &\leq \liminf_k \int_{\mathbb{R}^s} [f(x_k, z) - \eta]_+ \mu_n(dz) \end{aligned}$$

for all  $\eta \in \mathbb{R}$ . Taking the limes inferior with respect to  $n$  on both sides we obtain

$$\begin{aligned} \liminf_n \int_{\mathbb{R}^s} [f(x, z) - \eta]_+ \mu_n(dz) &\leq \liminf_n \liminf_k \int_{\mathbb{R}^s} [f(x_k, z) - \eta]_+ \mu_n(dz) \\ &\leq \liminf_n \int_{\mathbb{R}^s} [f(x_n, z) - \eta]_+ \mu_n(dz) \\ &\leq \int_{\mathbb{R}} [a - \eta]_+ \nu(da) \quad \forall \eta \in \mathbb{R}. \end{aligned}$$

Here the second inequality follows from passing to the diagonal sequence where  $n = k$ , and the third inequality follows from (15). Applying Lemma 2.1 with  $h(z) := [f(x, z) - \eta]_+$  implies

$$\int_{\mathbb{R}^s} [f(x, z) - \eta]_+ \mu(dz) \leq \liminf_n \int_{\mathbb{R}^s} [f(x_n, z) - \eta]_+ \mu_n(dz) \leq \int_{\mathbb{R}} [a - \eta]_+ \nu(da) \quad \forall \eta \in \mathbb{R}$$

and thus  $x \in C(\mu)$ . The proof is complete.  $\square$

**Remark 2.3** (About closedness of the sets  $C(\mu)$ .) By setting  $\mu_n$  identical to  $\mu$  for all  $n$ , Proposition 2.2 implies that  $C(\mu)$  is a closed subset of  $\mathbb{R}^m$  for all  $\mu \in \mathcal{P}_1(\mathbb{R}^s)$ .

**Remark 2.4** (About convexity of the sets  $C(\mu)$ .) Assume there are no integer variables in the second stage, i.e.,  $\Phi(t) := \min\{q^\top y : Wy = t, y \geq 0\}$  and  $X$  is convex. The convexity of  $\Phi$ , recall (10), then implies that for all  $x_1, x_2 \in X$  and all  $\lambda$  with  $0 \leq \lambda \leq 1$

$$\begin{aligned} [f(\lambda x_1 + (1 - \lambda)x_2, z) - \eta]_+ &\leq [\lambda(f(x_1, z) - \eta) + (1 - \lambda)(f(x_2, z) - \eta)]_+ \\ &\leq \lambda[f(x_1, z) - \eta]_+ + (1 - \lambda)[f(x_2, z) - \eta]_+. \end{aligned}$$

Together with (8) this yields the convexity of  $C(\mu)$  for all  $\mu \in \mathcal{P}_1(\mathbb{R}^s)$ .

**Remark 2.5** (About variable  $\nu$ .) In [10] the authors have studied the stability of first-order stochastic dominance constraints (involving generic random variables) when perturbing the underlying probability distributions for the data and the benchmark. When equipping the space  $\mathcal{P}_1(\mathbb{R})$  of benchmark measures  $\nu$  with weak convergence of probability measures and selecting the benchmarks from the subset  $\mathcal{P}_{\rho, R}(\mathbb{R}) \subseteq \mathcal{P}_1(\mathbb{R})$  of measures whose  $\rho$ -th moment is bounded above by  $R$  ( $\rho > 1, R > 0$  fixed), then  $\nu_n, \nu \in \mathcal{P}_{\rho, R}(\mathbb{R})$  and  $\nu_n \xrightarrow{w} \nu$  imply  $\int_{\mathbb{R}} [a - \eta]_+ \nu_n(da) \rightarrow \int_{\mathbb{R}} [a - \eta]_+ \nu(da)$ , see for instance [4], Theorem 5.4, p. 32. This enables straightforward extension of the proof of Proposition 2.2 to the multifunction  $\bar{C} : \mathcal{P}_1(\mathbb{R}^s) \times \mathcal{P}_{\rho, R}(\mathbb{R}) \rightarrow 2^{\mathbb{R}^m}$  where  $\bar{C}(\mu, \nu) := \{x \in \mathbb{R}^m : f(x, z) \succeq_2 a, x \in X\}$ .

**Remark 2.6** (About lower semicontinuity of the optimal value.) Closedness of the multifunction  $C$  is the key to proving lower semicontinuity of the optimal value function given by  $\varphi(\mu) := \inf\{g(x) : x \in C(\mu)\}$ . For instance, if  $X$  is nonempty and compact,  $g$  lower semicontinuous, and (A1)-(A3) are valid, then  $\varphi$  is lower semicontinuous at all  $\bar{\mu} \in \mathcal{P}_1(\mathbb{R}^s)$  for which the optimization problem defining  $\varphi(\bar{\mu})$  is solvable. The proof follows the lines of Berge's classical theory, see for instance [2] or [14].

### 3 Algorithm

For discrete probability distributions, the following proposition establishes an equivalence between (7) and a mixed-integer linear program.

**Proposition 3.1** Let  $z(\omega)$  and  $a(\omega)$  in (7) follow discrete distributions with realizations  $z_l$ ,  $l = 1, \dots, L$ , and  $a_k$ ,  $k = 1, \dots, K$ , as well as probabilities  $\pi_l$ ,  $l = 1, \dots, L$ , and  $p_k$ ,  $k = 1, \dots, K$ , respectively. Let further  $g(x) := g^\top x$  be linear. Assume (A1) and (A2). Then (7) is equivalent to the mixed-integer linear program

$$\min \left\{ \begin{array}{lll} g^\top x : & c^\top x + q^\top y_{lk} - a_k & \leq v_{lk} \quad \forall l \forall k \\ & Tx + Wy_{lk} & = z_l \quad \forall l \forall k \\ & \sum_{l=1}^L \pi_l v_{lk} & \leq \bar{a}_k \quad \forall k \\ & x \in X, y_{lk} \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}, \quad v_{lk} \geq 0 & \forall l \forall k \end{array} \right\} \quad (16)$$

where  $\bar{a}_k := \int_{\mathbb{R}} [a - a_k]_+ \nu(da)$ ,  $k = 1, \dots, K$ .

**Proof:** Recall from (8) that the constraint  $f(x, \omega) \succeq_2 a(\omega)$  is equivalent to

$$\int_{\mathbb{R}^s} [f(x, z) - \eta]_+ \mu(dz) \leq \int_{\mathbb{R}} [a - \eta]_+ \nu(da) \quad \forall \eta \in \mathbb{R}. \quad (17)$$

We first show that, for this to hold, validity for  $\eta = a_1, \dots, a_K$  is already sufficient. Denote for all  $\eta \in \mathbb{R}$

$$F(\eta) := \int_{\mathbb{R}^s} [f(x, z) - \eta]_+ \mu(dz)$$

and

$$A(\eta) := \int_{\mathbb{R}} [a - \eta]_+ \nu(da),$$

and assume that

$$F(a_k) \leq A(a_k), \quad k = 1, \dots, K. \quad (18)$$

Assume the  $a_k$  are arranged in ascending order. Since  $\nu$  is finite discrete, the function  $A$  is piecewise linear and convex. Moreover,  $A$  is linear on each of the intervals  $\eta \leq a_1, a_k \leq \eta \leq a_{k+1}$  ( $k = 1, \dots, K-1$ ), and  $a_K \leq \eta$ . We will check (17) for each of these intervals and start with  $\eta \leq a_1$ .

Pick  $\eta_o \leq \eta$  such that  $\eta_o \leq f(x, z_l)$  for all  $l = 1, \dots, L$ . Denoting  $L' := \{l \in \{1, \dots, L\} : f(x, z_l) < a_1\}$  we compute

$$\begin{aligned} F(\eta_o) - F(a_1) &= \sum_{l=1}^L \pi_l f(x, z_l) - \eta_o - \sum_{l \notin L'} \pi_l f(x, z_l) + \sum_{l \notin L'} \pi_l a_1 \\ &= \sum_{l \notin L'} \pi_l a_1 - \eta_o + \sum_{l \in L'} \pi_l f(x, z_l) \\ &\leq \sum_{l \notin L'} \pi_l a_1 - \eta_o + \sum_{l \in L'} \pi_l a_1 \\ &= a_1 - \eta_o. \end{aligned} \quad (19)$$

Moreover,  $A(\eta) = \sum_{k=1}^K p_k a_k - \eta$  for  $\eta \leq a_1$ , and

$$A(\eta_o) - A(a_1) = \sum_{k=1}^K p_k a_k - \eta_o - \sum_{k=1}^K p_k a_k + a_1 = a_1 - \eta_o. \quad (20)$$

For a suitable  $\lambda$  with  $0 \leq \lambda \leq 1$  it holds  $\eta = \lambda \eta_o + (1 - \lambda) a_1$ . The convexity of  $F$  then yields

$$\begin{aligned} F(\eta) &\leq \lambda F(\eta_o) + (1 - \lambda) F(a_1) \\ &= F(a_1) + \lambda (F(\eta_o) - F(a_1)) \\ &\leq F(a_1) + \lambda (a_1 - \eta_o) \\ &\leq A(a_1) + \lambda (a_1 - \eta_o) \\ &= A(a_1) + \lambda (A(\eta_o) - A(a_1)) \\ &= A(\lambda \eta_o + (1 - \lambda) a_1) \\ &= A(\eta). \end{aligned}$$

In the third row we have used (19), in the fourth (18), in the fifth (20), and then the linearity of  $A$  on the considered interval.

Now let  $a_k \leq \eta \leq a_{k+1}$  for some  $k = 1, \dots, K-1$ . The convexity of  $F$ , (18), and the linearity of  $A$  on this interval then provide for a suitable  $\lambda$  with  $0 \leq \lambda \leq 1$

$$F(\eta) \leq \lambda F(a_k) + (1-\lambda)F(a_{k+1}) \leq \lambda A(a_k) + (1-\lambda)A(a_{k+1}) = A(\lambda a_k + (1-\lambda)a_{k+1}) = A(\eta).$$

Finally, let  $a_K \leq \eta$ . Note that  $F$  is non-increasing and non-negative. Hence, together with (18),

$$0 \leq F(\eta) \leq F(a_K) \leq A(a_K) = 0.$$

Therefore,  $F(\eta) = 0 = A(\eta)$ . This proves the claim that (17) holds when it is valid for  $\eta = a_1, \dots, a_K$ . To establish the asserted equivalence we fix  $k$ , consider the sets

$$S_1 := \left\{ x \in X : \int_{\mathbb{R}^s} [f(x, z) - a_k]_+ \mu(dz) \leq \int_{\mathbb{R}} [a - a_k]_+ \nu(da) \right\}$$

and

$$S_2 := \left\{ x \in X : \begin{array}{l} \exists v_l \geq 0 \quad \exists y_l \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}, \quad l = 1, \dots, L, \\ \text{such that:} \\ c^\top x + q^\top y_l - a_k \leq v_l \\ Tx + Wy_l = z_l \\ \sum_{l=1}^L \pi_l v_l \leq \bar{a}_k \end{array} \right\}$$

and show that  $S_1 = S_2$ .

For  $S_1 \subseteq S_2$  let  $x \in S_1$  and denote  $I := \{l \in \{1, \dots, L\} : f(x, z_l) - a_k > 0\}$ . By the definition of  $S_1$  we have

$$\int_{\mathbb{R}^s} [f(x, z) - a_k]_+ \mu(dz) = \sum_{l \in I} \pi_l (f(x, z_l) - a_k) \leq \bar{a}_k.$$

Put  $v_l := f(x, z_l) - a_k$  for all  $l \in I$ , and  $v_l := 0$ , otherwise. This yields

$$\sum_{l=1}^L \pi_l v_l \leq \bar{a}_k.$$

For  $l \notin I$  it holds that  $f(x, z_l) - a_k \leq 0$ . The validity of (A1) and (A2) implies that the optimization problems behind the  $f(x, z_l)$  are solvable. Hence, for all  $l \notin I$ , there exist  $y_l \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}$  with

$$c^\top x + q^\top y_l - a_k \leq 0 = v_l \quad \text{and} \quad Tx + Wy_l = z_l.$$

For  $l \in I$ , choose  $y_l \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}$  such that  $q^\top y_l = \Phi(z_l - Tx)$  and  $Tx + Wy_l = z_l$ . Then

$$c^\top x + q^\top y_l - a_k = f(x, z_l) - a_k = v_l,$$

yielding  $x \in S_2$ .

For  $S_2 \subseteq S_1$  let  $x \in S_2$  and consider  $I := \{l \in \{1, \dots, L\} : v_l > 0\}$ . The definition of  $S_2$  implies that for  $l \notin I$  there exist  $y_l \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}$  fulfilling

$$c^\top x + q^\top y_l - a_k \leq 0 \quad \text{and} \quad Tx + Wy_l = z_l.$$

Therefore,  $f(x, z_l) - a_k \leq 0$  for all  $l \notin I$ . For  $l \in I$  there exist  $y_l \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}$  with

$$c^\top x + q^\top y_l - a_k \leq v_l \quad \text{and} \quad Tx + Wy_l = z_l.$$

Thus,  $f(x, z_l) - a_k \leq v_l$  for all  $l \in I$ . Now we obtain

$$\begin{aligned} \int_{\mathbb{R}^s} [f(x, z) - a_k]_+ \mu(dz) &= \sum_{l \in I} \pi_l [f(x, z_l) - a_k]_+ + \sum_{l \notin I} \pi_l [f(x, z_l) - a_k]_+ \\ &\leq \sum_{l \in I} \pi_l v_l + 0 \leq \sum_{l=1}^L \pi_l v_l \leq \bar{a}_k, \end{aligned}$$

so  $x \in S_1$ , and the proof is complete.  $\square$

Concerning the above proof we remark that the fact that, for finite probability spaces, second-order stochastic dominance reduces to a finite number of real-valued inequalities, cf. (18), has already been observed in [22]. The setting in [22] refers to a different class of random variables and stochastic dominance with preference of big outcomes, though.

Inspecting (16) we observe that the constraints

$$\sum_{l=1}^L \pi_l v_{lk} \leq \bar{a}_k \quad \forall k \quad (21)$$

are the only ones coupling explicitly second-stage variables, namely  $v_{lk}$ , across different scenarios  $l$ . An implicit such coupling, of course, is given by

$$\begin{aligned} c^\top x + q^\top y_{lk} - a_k &\leq v_{lk} & \forall l \forall k, \\ Tx + Wy_{lk} &= z_l & \forall l \forall k. \end{aligned}$$

One concludes that, without (21), problem (16) in principle were in  $L$ -shaped form, [30], a structure that has given rise to different decomposition algorithms for stochastic programs [6, 8, 18, 23, 25, 26].

Understanding (16) as an “expanded” representation of the nonconvex global minimization problem (7) we propose the following branch-and-bound algorithm for its solution. By  $\mathbf{P}$  we denote a list of problems, and  $\varphi_{LB}(P)$  is a lower bound for the optimal value of  $P \in \mathbf{P}$ . Moreover,  $\bar{\varphi}$  denotes the currently best upper bound to the optimal value of (16), and  $X(P)$  is the element in the partition of  $X$  belonging to  $P$ .

### Algorithm 3.2

STEP 1 (INITIALIZATION):

Let  $\mathbf{P} := \{(16)\}$  and  $\bar{\varphi} := +\infty$ .

STEP 2 (TERMINATION):

If  $\mathbf{P} = \emptyset$  then the  $\bar{x}$  that yielded  $\bar{\varphi} = g^\top \bar{x}$  is optimal.

STEP 3 (BOUNDING):

Select and delete a problem  $P$  from  $\mathbf{P}$ . Compute a lower bound  $\varphi_{LB}(P)$  and apply a feasibility heuristics to find a feasible point  $\bar{x}$  of  $P$ .

STEP 4 (PRUNING):

If  $\varphi_{LB}(P) = +\infty$  (infeasibility of a subproblem) or  $\varphi_{LB}(P) > \bar{\varphi}$  (inferiority of  $P$ ), then go to Step 2.

If  $\varphi_{LB}(P) = g^\top \bar{x}$  (optimality for  $P$ ), then check whether  $g^\top \bar{x} < \bar{\varphi}$ . If yes, then  $\bar{\varphi} := g^\top \bar{x}$ . Go to Step 2.

If  $g^\top \bar{x} < \bar{\varphi}$ , then  $\bar{\varphi} := g^\top \bar{x}$ .

STEP 5 (BRANCHING):

Create two new subproblems by partitioning the set  $X(P)$  by means of linear inequalities. Add these subproblems to  $\mathbf{P}$  and go to Step 2.

Of course, Algorithm 3.2 is of little value as long as the bounding in Step 3 is not specified. Let us start with lower bounding. The basic idea is to pass to a model in  $L$ -shaped form by means of relaxation. In view of the above discussion the obvious candidate for this relaxation is (21). Recall that there are as



many constraints in (21) as there are realizations of the benchmark  $a(\omega)$ . The latter originating from a subjective perception of risk, this number  $K$  often is quite small, say within some tens, compared with the generally far bigger number  $L$  of data scenarios  $z_l$ . Hence, Lagrangean relaxation of (21) will lead to a Lagrangean dual of tractable dimension.

For models in  $L$ -shaped form, two principal decomposition approaches can be taken, a Benders-type decomposition or a dual decomposition based on (Lagrangean) relaxation of nonanticipativity [6, 8, 18, 23, 25, 26, 30]. With integer variables in the second-stage, however, Benders decomposition leads to nonconvex master problems. Therefore, we pursue dual decomposition.

We relax nonanticipativity of  $x$  in (16) by introducing copies  $x_l, l = 1, \dots, L$ . One possibility now could be to regain nonanticipativity by Lagrangean relaxation of the identities  $x_1 = x_2 = \dots = x_L$ . This, however, would lead to a Lagrangean dual in dimension  $(L-1) \cdot \dim x$ , which quickly can become several tens or even hundreds of thousands. Therefore, we leave it at working with the copies  $x_l$  in our lower bounding scheme, striking a compromise between computational effort and quality of bounds.

With these presuppositions, and putting  $x = \sum_{l=1}^L \pi_l x_l$  we arrive at the following Lagrangean function

$$\begin{aligned} \mathcal{L}(x, v, \lambda) &= \sum_{l=1}^L \pi_l \cdot g^\top x_l + \sum_{k=1}^K \lambda_k \left( \sum_{l=1}^L \pi_l v_{lk} - \bar{a}_k \right) \\ &= \sum_{l=1}^L \pi_l \cdot g^\top x_l + \sum_{l=1}^L \sum_{k=1}^K \lambda_k \cdot (\pi_l v_{lk} - \pi_l \bar{a}_k) \\ &= \sum_{l=1}^L \mathcal{L}_l(x_l, v_l, \lambda) \end{aligned}$$

where

$$\mathcal{L}_l(x_l, v_l, \lambda) := \pi_l \cdot g^\top x_l + \pi_l \sum_{k=1}^K \lambda_k \cdot (v_{lk} - \bar{a}_k).$$

The Lagrangean dual reads

$$\max\{D(\lambda) : \lambda \in \mathbb{R}_+^K\}$$

where

$$D(\lambda) = \min \left\{ \mathcal{L}(x, v, \lambda) : \begin{array}{lll} c^\top x_l + q^\top y_{lk} - a_k & \leq & v_{lk} \quad \forall l \forall k \\ Tx_l + Wy_{lk} & = & z_l \quad \forall l \forall k \\ x_l \in X, y_{lk} \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}, v_{lk} \geq 0 & & \forall l \forall k \end{array} \right\}$$

This is where decomposition becomes effective. The optimization problem behind  $D(\lambda)$  is separable in  $l$ , and we obtain

$$D(\lambda) = \sum_{l=1}^L \min \left\{ \mathcal{L}_l(x_l, v_l, \lambda) : \begin{array}{lll} c^\top x_l + q^\top y_{lk} - a_k & \leq & v_{lk} \quad \forall k \\ Tx_l + Wy_{lk} & = & z_l \quad \forall k \\ x_l \in X, y_{lk} \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}, v_{lk} \geq 0 & & \forall k \end{array} \right\} \quad (22)$$

The function  $D(\cdot)$  is piecewise linear and concave. So bundle-trust methods for nonsmooth convex minimization can be employed for solving the Lagrangean dual, whose optimal value provides a lower bound for the optimal value of (16). In our numerical experiments we have used Christoph Helmberg's implementation of the spectral bundle method from [16].

As already mentioned there, upper bounding in Algorithm 3.2 is accomplished by a feasibility heuristics. This heuristics starts with  $x_l$ -parts  $\tilde{x}_l$  of optimal solutions to the single-scenario problems in (22) for optimal or nearly optimal  $\lambda$ .

### Algorithm 3.3

STEP 1:

Using  $\tilde{x}_l$ ,  $l = 1, \dots, L$ , pick a “reasonable candidate”  $\bar{x}$ , for instance one arising most frequently, or one with minimal  $\mathcal{L}_l(x_l, v_l, \lambda)$ , or average the  $\tilde{x}_l$ ,  $l = 1, \dots, L$ , and round to integers if necessary.

STEP 2:

Check whether the following problems are feasible for  $l = 1, \dots, L$ :

$$\min \left\{ g^\top \bar{x} : \begin{array}{ll} c^\top \bar{x} + q^\top y_{lk} - a_k & \leq v_{lk} \\ T\bar{x} + W y_{lk} & = z_l \\ y_{lk} \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}, & v_{lk} \geq 0, \quad k = 1, \dots, K \end{array} \right\} \quad (23)$$

If one of them fails to be feasible,  $\bar{x}$  cannot be feasible for (16), and the heuristics stops with assigning the formal upper bound  $+\infty$ . Otherwise, go to Step 3.

STEP 3:

Check whether the  $v_{lk}$  found in (23) fulfil

$$\sum_{l=1}^L \pi_l v_{lk} \leq \bar{a}_k \quad k = 1, \dots, K.$$

If yes, then a feasible solution to (16) is found. The heuristics stops with the upper bound  $g^\top \bar{x}$ . Otherwise, go to Step 4.

STEP 4:

Solve for each  $l = 1, \dots, L$ :

$$\min \left\{ \sum_{k=1}^K v_{lk} : \begin{array}{ll} c^\top \bar{x} + q^\top y_{lk} - a_k & \leq v_{lk} \\ T\bar{x} + W y_{lk} & = z_l \\ y_{lk} \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}, & v_{lk} \geq 0, \quad k = 1, \dots, K \end{array} \right\}$$

Go to Step 5.

STEP 5:

Repeat the test from Step 3 with the  $v_{lk}$  found in Step 4. If the test is positive then the heuristics stops with the upper bound  $g^\top \bar{x}$ . Otherwise, the heuristics stops without a feasible solution to (16) and assigns the formal upper bound  $+\infty$ .

## 4 Computations

In the following we report computational results for Algorithm 3.2 applied to test instances from power planning and Sudoku puzzling. The first group of instances refers to the optimal management of a dispersed generation (DG) system run by a power utility in Germany, see [15] for a detailed model description. The instances of the second group are inspired by [17].

### 4.1 Dispersed Generation System

The system contains five engine-based cogeneration stations which produce heat and power simultaneously and include altogether 18 generation units. Excessive heat is either stored or is exhausted (lost) via cooling devices. The electrical energy is fed into the distribution network. The system is completed by twelve wind turbines and one hydroelectric power plant.

Optimizing the operation of the system over some time horizon can be accomplished by a mixed-integer linear model that consists of about 17500 variables (9000 boolean, 8500 continuous) and 22000 constraints, if a horizon of 24 hours is split into quarter-hourly subintervals.

Taking into account data uncertainty - the consumers' demand of energy as well as the infeed from renewable resources and the fuel and power prices are known only in terms of probability distributions - the problem turns into a random optimization problem (2).

Finally, a two-stage stochastic optimization problem (3) arises, if we assume the data to be certain for the first four hours of the planning horizon. Results for a purely expectation-based model (6) and for mean-risk models with different risk measures are reported in [15, 27].

Dominance constraints are applied to this problem to minimize abrasion of the generation units over all possible generation policies that dominate a given cost benchmark to second order. Therefore the new objective function  $g(x)$  in (7) is chosen as the sum over all variables indicating start-ups of the units in the first four hours. The benchmark  $a$  is provided by an optimal solution  $\hat{x}$  of the expectation-based model (6). The  $f(\hat{x}, \omega)$  are clustered around some heuristically chosen benchmark values and each benchmark value is assigned the sum of the probabilities of the members in its cluster. Further test instances are obtained by successively increasing the benchmark values, with probabilities fixed.

Below, computational results for  $K = 4$  benchmark scenarios and  $L = 20$  up to  $L = 50$  scenarios for heat and power demand are reported. Dimensions of the arising deterministic equivalents are shown in Table 1.

Number of	20 scenarios	30 scenarios	50 scenarios
Boolean variables	596719	894319	1489519
continuous variables	564693	846333	1409613
constraints	1481568	2220488	3698328

Table 1: Dimensions of mixed-integer linear programming equivalents

In Tables 2 - 4 we compare results for the deterministic equivalents gained from the standard mixed-integer solver Cplex ([9]) to results computed with Algorithm 3.2 derived in Section 3, called ddsip.vSD here. Computations were done on a Linux-PC with a 3.2GHz pentium processor and 2GB ram. As stopping criterion we used a timelimit of eight hours.

From instance 1 to instance 5 the benchmark costs increase successively which makes the dominance constraints easier to fulfil. As one would expect, this leads to a decrease in the number of start-ups of the generation units. This is reported in the column 'Upper Bound', where the objective value of the current best solution is displayed. The corresponding lower bounds ('Lower Bound') are given as well.

In all test instances, ddsip.vSD reaches the first feasible solution faster than Cplex does. The time ddsip.vSD needs to find this first feasible solution is thus reported. Furthermore, the points in time are given, where ddsip.vSD and Cplex, respectively, solve the test instances to optimality. For Cplex, this always turned out to be the time when the first feasible solution was found. Finally, if a computation was not finished within eight hours, its status at expiry of the timelimit is reported.

Instance	Benchmarks		Time (sec.)	Cplex		ddsip.vSD	
	Probability	Benchmark Value		Upper Bound	Lower Bound	Upper Bound	Lower Bound
1	0.105	2895000	455.36	-	29	29	9
	0.1	4851000	1602.92	-	29	29	29
	0.69	7789000	3694.92	29	29	29	29
	0.105	10728000					
2	0.105	2900000	464.26	-	27	27	9
	0.1	4860000	2501.01	-	27	27	27
	0.69	7800000	4661.86	27	27	27	27
	0.105	10740000					
3	0.105	3000000	362.91	-	18	18	9
	0.1	5000000	3101.28	-	18	18	18
	0.69	8000000	3952.93	18	18	18	18
	0.105	11000000					
4	0.105	3500000	386.32	-	11	11	9
	0.1	5500000	1197.34	-	11	11	11
	0.69	8500000	3760.04	11	11	11	11
	0.105	11500000					
5	0.105	4000000	430.61	-	8	8	8
	0.1	6000000	3035.52	8	8	8	8
	0.69	9000000					
	0.105	12000000					

Table 2: Results for instances with 20 data scenarios and 4 benchmark scenarios

For one of the test instances with 30 scenarios Cplex stopped before reaching a feasible solution, because

the available memory was exceeded (marked by 'out of mem.'). In this case the lower bound at break off is reported.

With 50 data scenarios the deterministic equivalents get so large, that the available memory is insufficient to build up the model (lp-) file used by Cplex, thus preventing optimization with Cplex for these instances. Hence, in the last table only best values and lower bounds found by ddsip.vSD are presented.

Instance	Benchmarks		Time (sec.)	Cplex		ddsip.vSD	
	Probability	Benchmark Value		Upper Bound	Lower Bound	Upper Bound	Lower Bound
1	0.085	2895000	697.21	–	29	29	9
	0.14	4851000	2471.63	–	29	29	29
	0.635	7789000	7520.07	29	29	29	29
	0.14	10728000					
2	0.085	2900000	702.31	–	27	27	9
	0.14	4860000	3635.25	–	27	27	27
	0.635	7800000	14905.68	–	27 out of mem.	27	27
	0.14	10740000					
3	0.085	3000000	666.31	–	18	18	9
	0.14	5000000	3907.92	–	18	18	18
	0.635	8000000	7181.68	18	18	18	18
	0.14	11000000					
4	0.085	3500000	500.05	–	11	11	9
	0.14	5500000	1404.96	–	11	11	11
	0.635	8500000	6559.52	11	11	11	11
	0.14	11500000					
5	0.085	4000000	474.68	–	8	8	8
	0.14	6000000	6076.40	8	8	8	8
	0.635	9000000					
	0.14	12000000					

Table 3: Results for instances with 30 data scenarios and 4 benchmark scenarios

Instance	Benchmarks		Time (sec.)	Cplex		ddsip.vSD	
	Probability	Benchmark Value		Upper Bound	Lower Bound	Upper Bound	Lower Bound
1	0.09	2895000	1084.68	–	–	29	9
	0.135	4851000	3747.69	–	–	29	29
	0.67	7789000					
	0.105	10728000					
2	0.09	2900000	1125.39	–	–	27	9
	0.135	4860000	5857.67	–	–	27	27
	0.67	7800000					
	0.105	10740000					
3	0.09	3000000	1041.15	–	–	18	9
	0.135	5000000	6126.89	–	–	18	18
	0.67	8000000					
	0.105	11000000					
4	0.09	3500000	1026.21	–	–	11	9
	0.135	5500000	2872.83	–	–	11	11
	0.67	8500000					
	0.105	11500000					
5	0.09	4000000	1096.69	–	–	8	8
	0.135	6000000					
	0.67	9000000					
	0.105	12000000					

Table 4: Results for instances with 50 data scenarios and 4 benchmark scenarios

Our computations show that the decomposition method solves all test instances to the optimum in less computing time than Cplex does. It always reaches a first feasible solution very fast, seven up to twenty times faster than Cplex, which seems to be the decisive difficulty for the standard solver.

Especially in the computations with 30 and 50 scenarios the superiority of the decomposition method over general-purpose solvers becomes apparent. For one instance with 30 scenarios Cplex cannot provide a feasible solution, for all instances with 50 scenarios even no lower bound. On the other hand, ddsip.vSD solves all these problems to optimality.

## 4.2 Sudoku

Sudoku is a popular logic game, which is played over a  $9 \times 9$  grid, canonically divided into nine  $3 \times 3$  sub grids. Sudoku begins with some of the grid cells already filled with numbers. The task of a Sudoku player is to fill the remaining empty cells with numbers between 1 and 9 (one number only in each cell), such that each number occurs only once in each row, each column and each of the nine sub blocks. The

Sudoku rules can easily be represented with 729 Boolean variables and a system of linear inequalities (cf. [17]).

A two-stage random mixed-integer linear program (2) arises in the following way: The entries on the main diagonal are chosen as first stage decisions. A scenario is formed by a single Sudoku puzzle with a small number of prescribed entries and the property that a solution with joint elements on the main diagonal exists. The objective is to minimize the sum of the elements of the secondary diagonal (north-east to south-west).

To arrive at a dominance constrained model (7), we choose the objective  $g(x) = g^\top x$  as the sum of the elements on the main diagonal. Benchmark scenarios were derived by clustering  $f(\hat{x}, \omega)$ , where  $\hat{x}$  denotes an optimal solution to the expectation model (6). In this way we minimize the sum of the main diagonal elements such that the corresponding member of (5) stochastically dominates the specified benchmark random variable to second order.

We report results with  $K = 1$  up to 5 benchmark scenarios and  $L = 10$  up to 100 scenarios. Deterministic equivalents according to Proposition 3.1 again become pretty large-scale. Table 5 shows dimensions for  $K = 5$  and some  $L$ .

Number of	10 scenarios	20 scenarios	50 scenarios	100 scenarios
boolean variables	36450	72900	182250	364500
general integer variables	9	9	9	9
continuous variables	50	100	250	500
constraints	20755	41505	103755	207505

Table 5: Dimensions of mixed-integer linear programming equivalents

Table 6 summarizes our computations for the Sudoku instances. Again, a Linux-PC with a 3.2GHz pentium processor and 2GB ram was used. The time limit was set to ten hours.

The first two columns list the numbers  $K$  of benchmark and  $L$  of data scenarios. The remaining columns list lower and upper bounds obtained when applying Cplex [9] and our implementation ddsip.vSD of Algorithm 3.2. Time entries deviating from the limit of 10h indicate that the instance was solved to optimality within this span. It can be seen that ddsip.vSD was able to solve all instances to optimality within the horizon of 10 hours, except the 100 scenario, 5 references problem, while Cplex did not find a feasible point in most cases.

Number of Bench- mark Scenarios	Number of Data Scenarios	Cplex			ddsip.vSD		
		Upper Bound	Lower Bound	Time ( <i>min</i> )	Upper Bound	Lower Bound	Time ( <i>min</i> )
1	10	-	45	600	45	45	4
	20	45	45	0.1	45	45	5
	50	45	45	0.4	45	45	5
	100	-	45	600	45	45	9
2	10	45	45	0.4	45	45	33
	20	45	45	0.6	45	45	22
	50	45	45	2.6	45	45	37
	100	-	45	600	45	45	67
3	10	-	45	600	45	45	94
	20	-	45	600	45	45	168
	50	-	45	600	45	45	96
	100	-	45	600	45	45	151
4	10	-	45	600	45	45	156
	20	-	45	600	45	45	260
	50	-	45	600	45	45	198
	100	-	45	600	45	45	201
To be continued on the next page							

Number of Bench- mark Scenarios	Number of Data Scenarios	Cplex			ddsip.vSD		
		Upper Bound	Lower Bound	Time ( <i>min</i> )	Upper Bound	Lower Bound	Time ( <i>min</i> )
5	10	-	45	600	45	45	413
	20	-	45	600	45	45	480
	50	-	45	600	45	45	229
	100	-	45	600	45	43	600

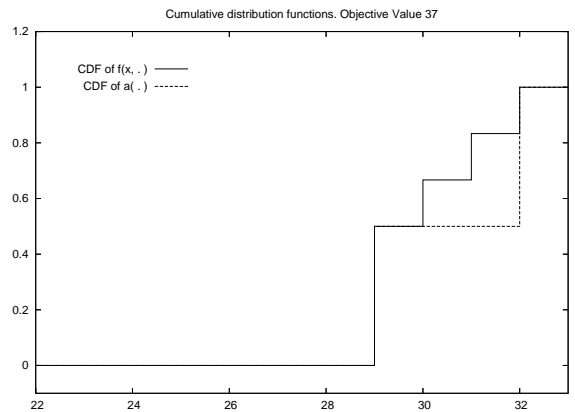
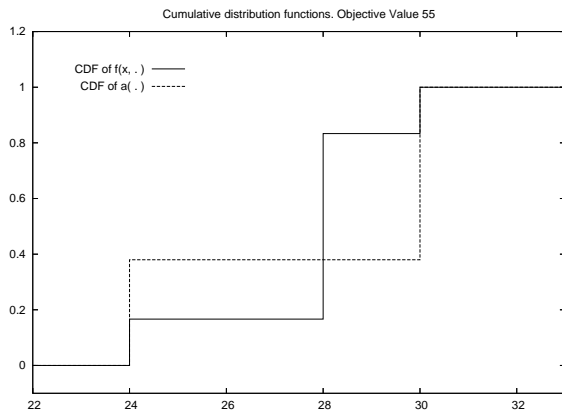
Table 6: Results for Sudoku instances

As an illustration of possible tradeoffs between the dominance model (7) and the expectation model (6) we did some computations with a fixed number of 6 scenarios and different (uniformly distributed) random variables  $a$ . The results of these experiments are shown in Table 7, where the first column contains the possible values of the random variable  $f(x^*, \omega)$  (also uniformly distributed) to the optimal  $x^*$  found by solving the dominance constrained model. The second column lists the different benchmark values (realizations of  $a$ ). It can be observed that the looser the benchmark profile is chosen the better is the objective value of the dominance model. On the other hand one has to cope with a higher expected value.

The figures below show the cumulative distribution functions (cdf) of  $a(\cdot)$  and  $f(x^*, \cdot)$  for the first and the last instances of Table 7. While in the first case  $f$  stochastically dominates  $a$  only to second-order, the cdf of  $f$  is greater or equal to that of  $a$  in the second case. Thus even dominance to first-order holds for this instance.

Values of $f(x^*, \omega)$	Benchmark	Objective Value (Dominance Model)	$E[f(x^*, \omega)]$
24, 28, 28, 28, 28, 30	24, 30	55	27.6667
27, 27, 28, 28, 28, 29	27, 29	44	27.8333
27, 28, 28, 28, 28, 30	26.5, 30	40	28.1667
29, 29, 29, 30, 31, 31	29, 31	38	29.8333
29, 29, 29, 30, 31, 32	29, 32	37	30

Table 7: Results for different benchmarks



**Acknowledgement.** We wish to thank Christoph Helmberg (Technical University of Chemnitz) for giving us access to the implementation of his spectral bundle method. Further thanks are due to Frederike Neise (University of Duisburg-Essen) for fruitful discussions and her support in computational testing. Parts of this paper were written while the third author was visiting the Centro de Modelamiento Matemático, Universidad de Chile, Santiago. Partial funding for this research was provided by the German

Federal Ministry of Education and Research (BMBF) under grant 03-SCNIVG. We thank these institutions for their support.

## References

- [1] Ahmed, S.: Convexity and decomposition of mean-risk stochastic programs, *Mathematical Programming* 106 (2006), 433 - 446.
- [2] Bank, B.; Guddat, J.; Klatte, D.; Kummer, B.; Tammer, K.: *Non-linear Parametric Optimization*, Akademie-Verlag, Berlin, 1982.
- [3] Bank, B.; Mandel, R.: *Parametric Integer Optimization*, Akademie-Verlag, Berlin, 1988.
- [4] Billingsley, P.: *Convergence of Probability Measures*, Wiley, New York, 1968.
- [5] Billingsley, P.: *Probability and Measure*, Wiley, New York, 1986.
- [6] Birge, J.R.; Louveaux, F.: *Introduction to Stochastic Programming*, Springer, New York, 1997.
- [7] Blair, C.E.; Jeroslow, R.G.: The value function of a mixed integer program: I. *Discrete Mathematics* 19 (1977), 121 - 138.
- [8] Carøe, C.C.; Schultz, R.: Dual decomposition in stochastic integer programming. *Operations Research Letters* 24 (1999), 37 - 45.
- [9] CPLEX Callable Library 9.1.3, ILOG (2005).
- [10] Dentcheva, D.; Henrion, R.; Ruszczyński, A.: Stability and sensitivity of optimization problems with first order stochastic dominance constraints, *SIAM Journal on Optimization* (submitted).
- [11] Dentcheva, D.; Ruszczyński, A.: Stochastic optimization with dominance constraints, *SIAM Journal on Optimization* 14 (2003), 548 - 566.
- [12] Dentcheva, D.; Ruszczyński, A.: Optimality and duality theory for stochastic optimization with nonlinear dominance constraints, *Mathematical Programming* 99 (2004), 329 - 350.
- [13] Eichhorn, A.; Römisch, W.: Polyhedral risk measures in stochastic programming, *SIAM Journal on Optimization* 16 (2005), 69 - 95.
- [14] Gollmer, R.; Neise, F.; Schultz, R.: Stochastic programs with first-order dominance constraints induced by mixed-integer linear recourse, Preprint 641-2006, Department of Mathematics, University of Duisburg-Essen, 2006.
- [15] Handschin, E.; Neise, F.; Neumann, H.; Schultz, R.: Optimal operation of dispersed generation under uncertainty using mathematical programming, *International Journal of Electrical Power & Energy Systems* 28 (2006), 618-626.
- [16] Helmberg, C.; Kiwiel, K.C.: A spectral bundle method with bounds, *Mathematical Programming* 93 (2002), 173-194.
- [17] Kaibel, V.; Koch, T.: Mathematik für den Volkssport, *DMV-Mitteilungen* 14 (2006), 93 - 96.
- [18] Kall, P.; Wallace, S.W.: *Stochastic Programming*, Wiley, Chichester, 1994.
- [19] Kristoffersen, T.K.: Deviation measures in two-stage stochastic linear programming, *Mathematical Methods of Operations Research* 62 (2006), 255 - 274.
- [20] Märkert, A.; Schultz, R.: On deviation measures in stochastic integer programming, *Operations Research Letters* 33 (2005), 441-449.
- [21] Müller, A.; Stoyan, D.: *Comparison Methods for Stochastic Models and Risks*, Wiley, Chichester, 2002.

- [22] Noyan, N.; Rudolf, G.; Ruszczyński, A.: Relaxations of linear programming problems with first order stochastic dominance constraints, *Operations Research Letters* 34 (2006), 653 - 659.
- [23] Prékopa, A.: *Stochastic Programming*, Kluwer, Dordrecht, 1995.
- [24] Römisch, W.: Stability of Stochastic Programming Problems, in [26], 483-554.
- [25] Ruszczyński, A.: Decomposition methods in stochastic programming, *Mathematical Programming* 79 (1997), 333–353.
- [26] Ruszczyński, A.; Shapiro, A. (eds.): *Handbooks in Operations Research and Management Science*, 10: *Stochastic Programming*, Elsevier, Amsterdam, 2003.
- [27] Schultz, R.; Neise, F.: Algorithms for mean-risk stochastic integer programs in energy, Preprint 615-2005, Department of Mathematics, University of Duisburg-Essen, Investigación Operacional (to appear).
- [28] Schultz, R.; Tiedemann, S.: Risk Aversion via Excess Probabilities in Stochastic Programs with Mixed-Integer Recourse. *SIAM Journal on Optimization* 14 (2003), 115–138.
- [29] Schultz, R.; Tiedemann, S.: Conditional value-at-risk in stochastic programs with mixed-integer recourse, *Mathematical Programming* 105 (2006), 365–386.
- [30] Van Slyke, R.M.; Wets, R. J-B: *L-shaped linear programs with applications to optimal control and stochastic linear programming*, *SIAM Journal on Applied Mathematics* 17 (1969), 638–663.